

Collision Avoidance in Multi-Hop Ad Hoc Networks *

Yu Wang J.J. Garcia-Luna-Aceves
Department of Computer Engineering
University of California
Santa Cruz, CA 95064, U.S.A.
Email: {ywang, jj}@cse.ucsc.edu

Abstract

Collision avoidance is very important in contention-based medium access control protocols for multi-hop ad hoc networks due to the adverse effects of hidden terminals. Four-way sender-initiated schemes are the most popular collision-avoidance schemes to date. Although there has been considerable work on the performance evaluation of these schemes, most analytical work is confined to single-hop ad hoc networks or networks with very few hidden terminals. In this paper, we use a simple analytical model to derive the saturation throughput of collision avoidance protocols in multi-hop ad hoc networks with nodes randomly placed according to a two-dimensional Poisson distribution, which to our knowledge has not been investigated sufficiently before. We show that the sender-initiated collision-avoidance scheme achieves much higher throughput than the idealized carrier sense multiple access scheme with an ideal separate channel for acknowledgments. More importantly, we show that the collision avoidance scheme can accommodate much fewer competing nodes within a region in a network infested with hidden terminals than in a fully-connected network, if reasonable throughput is to be maintained. This shows that the scalability problem of contention-based collision-avoidance protocols looms much earlier than people might expect. Simulation experiments of the popular IEEE 802.11 MAC protocol validate the predictions made in the analysis.

Keywords

Collision avoidance, MAC, ad hoc networks, IEEE 802.11, analytical modeling, simulation evaluation

1 Introduction

Collision avoidance in contention-based medium access control (MAC) protocols for multi-hop ad hoc networks is very important, as simple MAC protocols such as carrier sense multiple access (CSMA) cannot combat the “hidden terminal” problem and performance can degrade to that of the ALOHA protocol in ad hoc networks [14].

Many collision-avoidance protocols [2, 11] have been proposed and the most popular collision avoidance scheme today consists of a sender-initiated four-way handshake in which the transmission of a data packet and its acknowledgment is preceded by request-to-send (RTS) and clear-to-send (CTS) packets between a pair of sending and receiving nodes.¹ Other nodes that overhear RTS or CTS packets will defer their access to the channel to avoid collisions. For the sake of simplicity, it can be also called RTS/CTS-based scheme. Among all these proposed collision-avoidance protocols, the IEEE 802.11 distributed foundation wireless medium access control (DFWMAC) protocol [11] is very popular in the performance studies of routing protocols for ad hoc networks, even though it was originally intended for wireless LANs with no or very few hidden terminals. Though there has been considerable work on the performance evaluation of IEEE 802.11 and similar protocols [3–8, 17, 18], most of the analytical models are largely confined to single-hop networks [3, 5, 6] or cases when the number of hidden terminals is very small [7, 8]. We deem it very important to investigate the performance of the four-way sender-initiated collision avoidance scheme with a truly multi-hop network model as potential interference from hidden nodes always exists, which is a salient characteristic of multi-hop ad hoc networks.

In this paper, we adopt a simple multi-hop network model to derive the saturation throughput of a sender-

*This work was supported in part by the US Air Force/OSR under Grant No. F49620-00-1-0330.

¹This is in contrast to receiver-initiated MAC schemes (e.g. [10]) in which the collision avoidance handshake is started by receivers.

Report Documentation Page			Form Approved OMB No. 0704-0188		
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE 2002		2. REPORT TYPE		3. DATES COVERED 00-00-2002 to 00-00-2002	
4. TITLE AND SUBTITLE Collision Avoidance in Multi-Hop Ad Hoc Networks			5a. CONTRACT NUMBER		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S)			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of California at Santa Cruz, Department of Computer Engineering, Santa Cruz, CA, 95064			8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSOR/MONITOR'S ACRONYM(S)		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)		
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES 10	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

initiated collision avoidance scheme, in which nodes are randomly placed on a plane according to two-dimensional Poisson distribution with density λ . Varying λ has the effect of changing the congestion level within a region as well as the number of hidden terminals. The adoption of a Poisson distribution is due to its tractability for analysis and ability to model multi-hop networks. With this model, we present the analysis of the basic four-way sender-initiated collision-avoidance scheme in Section 2, which to our knowledge is the first analytical modeling of collision avoidance in multi-hop networks.

In Section 3, we present numerical results from our analysis. We compare the performance of the sender-initiated collision avoidance scheme against the idealized non-persistent CSMA protocol in which a secondary channel is assumed to send acknowledgments in zero time and without collisions [14, 19], as the latter is the only protocol whose analysis for multi-hop ad hoc networks is available for comparison to date. It is shown that the RTS/CTS scheme can achieve far better throughput than the CSMA protocol, even when the overhead due to the RTS/CTS exchange is high. The results illustrate the importance of enforcing collision avoidance in the RTS/CTS handshake. It is also shown that, the aggregate throughput of sender-initiated collision avoidance drops faster than that in a fully-connected network (or single-hop networks) when the number of competing nodes within a region increases. That is, collision avoidance becomes more and more ineffective for a relatively crowded region with hidden terminals.

In Section 4 we present simulation experiments of the popular IEEE 802.11 MAC protocol to validate the predictions in the analysis. Section 5 concludes this paper with possible ways to improve the performance of collision avoidance protocols in ad hoc networks.

2 Approximate Analysis

In this Section, we derive the approximate throughput of a perfect collision avoidance protocol in the network model we investigate. Here *perfect collision avoidance* means that nodes can accurately sense the channel busy or idle, and that the RTS/CTS scheme can avoid the transmission of data packets that collide with other packets at the receivers. This has been shown to be achievable in the floor acquisition multiple access (FAMA) protocol [9] by enforcing longer CTS packets to serve as “busy tone” packet to silence hidden terminals as well as various waiting times.

In the network model, nodes are two-dimensionally Poisson distributed with density λ , i.e., the probability $p(i, S)$ of finding i nodes in an area of S is given by:

$$p(i, S) = \frac{(\lambda S)^i}{i!} e^{-\lambda S}.$$

Assume that each node has the same transmission and receiving range of R , and denote by N the average number of nodes within a circular region of radius R ; therefore, we have $N = \lambda \pi R^2$.

To simplify our analysis, we assume that nodes operate in time-slotted mode, which is reasonable when the maximum propagation delay is much smaller than packet transmission time. In such a case, the performance of the slotted system will be much the same as the system that is not time-slotted [14]. The length of each time slot is denoted by τ , which includes propagation delay as well as the overhead such as the transmit-to-receive turn-around time, carrier sensing delay and processing time. Hence τ represents the time required for all the nodes within the transmission range to know the event that occurred τ seconds ago. The transmission times of RTS, CTS, data, and ACK packets are normalized with regard to τ and are denoted by l_{rts} , l_{cts} , l_{data} , and l_{ack} , respectively. Thus τ simply equals 1. For the sake of simplicity, we also assume that all packet transmission times are multiples of the length of a time-slot.

We derive the protocol’s throughput based on the heavy-traffic assumption, i.e., a node always has a packet in its buffer to be sent and the destination is chosen randomly from one of its neighbors. This is a fair assumption for ad hoc networks in which nodes are sending data and signaling packets continually. We also assume that a node is ready to transmit with probability p and not ready with probability $1 - p$. Here, p is a protocol-specific parameter that is slot independent. At the level of individual nodes, the probability of being ready to transmit may vary from time slot to slot, depending on the current states of both the channel and the node. However, because we are interested in deriving the average performance metrics instead of instantaneous or short-term metrics, the assumption of a fixed probability p may be considered as an averaged quantity that can still reasonably approximate the factual burstiness from a long-term point of view. In fact, this assumption is necessary to make the theoretical modeling tractable and has been extensively applied before by some researchers [5, 13, 19]. For example, the model was used by Takagi and Kleinrock [13] to derive the optimal transmission range of a node in a multi-hop wireless network, and was used subsequently by Wu and Varshney [19] to derive the throughput of non-persistent CSMA and some variants of busy tone multiple access (BTMA) protocols [14], whose results will be compared later in this paper.

It should also be noted that, even when a node is ready to transmit, it may transmit or not in the slot, depending on the collision avoidance and resolution schemes being used, as well as the channel’s current state. Thus, we are more interested in the probability that a node transmits in a time slot, which is denoted by p' . Similar to above reasoning, we also assume that p' is independent at any time slot to make

the analysis tractable. Given this simplification, p' can be defined to be:

$$\begin{aligned} p' &= p \cdot \text{Prob.}\{\text{Channel is sensed idle in a slot}\} \\ &\approx p \cdot \Pi_I \end{aligned}$$

where Π_I is the limiting probability that the channel is in idle state, which we derive subsequently.

We are not interested in the exact relationship between p and p' , and it is enough to obtain the range of values that p' can take, because the throughput of these protocols is mostly influenced by p' . To derive the rough relationship between p and p' , we set up a channel model which includes two key simplifying assumptions.

First, we model the channel as a circular region in which there are some nodes. The nodes within the region can communicate with each other while they have weak interactions with nodes outside the region. *Weak interaction* means that the decision of inner nodes to transmit, defer and back off is almost unaffected by that of outer nodes and vice versa. Considering that nodes do not exchange status information explicitly (e.g. either defer due to collision avoidance or back off due to collision resolution), this assumption is reasonable and helps to simplify the model considerably. Thus the channel's status is only decided by the successful and failed transmissions within the region.

Second, we still consider the failed handshakes initiated by nodes within the region to outside nodes, because this has a direct effect on the channel's usability for other nodes within the region. Though the radius of the circular region R' is unknown, it falls between $R/2$ and $2R$. This follows from noting that the maximal radius of a circular region in which all nodes are guaranteed to hear one another equals $R' = R/2$, and all the direct neighbors and hidden nodes are included into the region when $R' = 2R$. Thus we can write $R' = \alpha R$ where $0.5 \leq \alpha \leq 2$, and α needs to be estimated.

With the above assumptions, the channel can be modeled by a four-state Markov chain illustrated in Figure 1. The significance of the states of this Markov chain is the following:

- *Idle* is the state when the channel around node x is sensed idle, and obviously its duration is τ .
- *Long* is the state when a successful four-way handshake is done. For simplicity, we assume that the channel is in effect busy for the duration of the whole handshake, thus the busy time T_{long} is

$$\begin{aligned} T_{long} &= l_{rts} + \tau + l_{cts} + \tau + l_{data} + \tau \\ &\quad + l_{ack} + \tau \\ &= l_{rts} + l_{cts} + l_{data} + l_{ack} + 4\tau. \end{aligned}$$

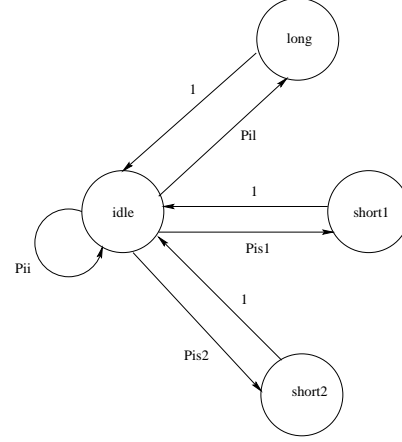


Figure 1. Markov chain model for the channel around a node

- *Short1* is the state when multiple nodes around the channel transmit RTS packets during the same time slot and their transmissions collide. The busy time of the channel T_{short1} is therefore

$$T_{short1} = l_{rts} + \tau.$$

- *Short2* is the state when one node around the channel initiates a failed handshake with a node outside the region. Even though a CTS packet may not be sent due to the collision of the sending node's RTS packet with other packets originated from nodes outside the region or due to the deferring of the receiving node to other nodes, those nodes overhearing the RTS as well as the sending node do not know if the handshake is successfully continued, until the time required for receiving a CTS packet elapses. Therefore the channel is in effect busy, i.e., unusable for all the nodes sharing the channel, for the time stated below:

$$\begin{aligned} T_{short2} &= l_{rts} + \tau + l_{cts} + \tau \\ &= l_{rts} + l_{cts} + 2\tau. \end{aligned}$$

Now we proceed to calculate the transition probabilities of the Markov chain.

In usual collision avoidance schemes, no node is allowed to transmit immediately after the channel becomes idle, thus the transition probabilities from *long* to *idle*, from *short1* to *idle* and from *short2* to *idle* are all one.

According to the Poisson distribution of the nodes, the probability of having i nodes within the receiving range R of x is $e^{-N} N^i / i!$, where $N = \lambda \pi R^2$. Therefore, the mean number of nodes that belong to the shared channel is $M = \lambda \pi R'^2 = \alpha^2 N$. Assuming that each node transmits

independently, the probability that none of them transmits is $(1 - p')^i$ where $(1 - p')$ is the probability that a node does not transmit in a time slot. Because the transition probability P_{ii} from *idle* to *idle* is the probability that none of the neighboring nodes of x transmits in this slot, P_{ii} is given by

$$\begin{aligned} P_{ii} &= \sum_{i=0}^{\infty} (1 - p')^i \frac{M^i}{i!} e^{-M} \\ &= e^{-p'M}. \end{aligned}$$

We average the probabilities over the number of interfering nodes in a region because of two reasons. First, it is much more tractable than the approach that conditions on the number of nodes, calculates the desired quantities and then uses the Poisson distribution to obtain the average. Second, in our simulation experiments, we also fix the number of competing nodes in a region (which is N) and then vary the location of the nodes to approximate the Poisson distribution, which is configurationally closer to our analytical model; the alternative would be to generate 2, 3, 4, ... nodes within one region, get the throughput for the individual configuration and then calculate the average, which is not practical.

Next we need to calculate the transition probability P_{il} from *idle* to *long*. If there are i nodes around node x , for such a transition to happen, one and only one node should be able to complete one successful four-way handshake while other nodes do not transmit. Let p_s denote the probability that a node begins a successful four-way handshake at each slot, we can then calculate P_{il} as follows:

$$\begin{aligned} P_{il} &= \sum_{i=1}^{\infty} i p_s (1 - p')^{i-1} \frac{M^i}{i!} e^{-M} \\ &= p_s M e^{-p'M}. \end{aligned}$$

To obtain the above result, we use the fact that the distribution of the number of nodes within R' does not depend on the existence of node x , because of the memoryless property of Poisson distribution. Up to this point, p_s is still an unknown quantity that we derive subsequently.

The transition probability from *idle* to *short1* is the probability that more than one node transmit RTS packets in the same slot; therefore, P_{is1} can be calculated as follows:

$$\begin{aligned} P_{is1} &= \sum_{i=2}^{\infty} [1 - (1 - p')^i - i p' (1 - p')^{i-1}] \cdot \frac{M^i}{i!} e^{-M} \\ &= 1 - (1 + M p') e^{-p'M}. \end{aligned}$$

Having calculated P_{ii} , P_{il} and P_{is1} , we can calculate P_{is2} , the transition probability from *idle* to *short2*, which is equal to $1 - P_{ii} - P_{il} - P_{is1}$. Let π_i , π_l , π_{s1} and π_{s2} denote the steady-state probabilities of states *idle*, *long*, *short1* and

short2 respectively. From Figure 1, we have

$$\begin{aligned} \pi_i P_{ii} + \pi_l + \pi_{s1} + \pi_{s2} &= \pi_i \\ \pi_i P_{ii} + 1 - \pi_i &= \pi_i \\ \pi_i &= \frac{1}{2 - P_{ii}} = \frac{1}{2 - e^{-p'M}}. \end{aligned}$$

The limiting probability Π_I , i.e., the long run probability that the channel around node x is found idle, can be obtained by:

$$\begin{aligned} \Pi_I &= \frac{\pi_i T_{idle}}{\pi_i T_{idle} + \pi_l T_{long} + \pi_{s1} T_{short1} + \pi_{s2} T_{short2}} \\ &= \frac{T_{idle}}{T_{idle} + P_{il} T_{long} + P_{is1} T_{short1} + P_{is2} T_{short2}}. \end{aligned}$$

($\pi_i P_{il} = \pi_l$, $\pi_i P_{is1} = \pi_{s1}$ and $\pi_i P_{is2} = \pi_{s2}$)

The relationship between p' and p is then:

$$p' = \frac{p T_{idle}}{T_{idle} + P_{il} T_{long} + P_{is1} T_{short1} + P_{is2} T_{short2}}. \quad (1)$$

In the above equation, the probability that a node x starts successfully a four-way handshake in a time slot, p_s , is yet to be determined.

We model the states of a node x by a three-state Markov chain shown in Figure 2. In Figure 2, *wait* is the state when the node defers for other nodes or backs off, *succeed* is the state when the node can complete a successful four-way handshake with other nodes, and *fail* is the state when the node initiates an unsuccessful handshake. For simplicity, we regard *succeed* and *fail* as the states when two different kinds of *virtual* packets are transmitted and their lengths are:

$$\begin{aligned} T_{succeed} &= T_{long} \\ &= l_{rts} + l_{cts} + l_{data} + l_{ack} + 4\tau \\ T_{fail} &= T_{short2} \\ &= l_{rts} + l_{cts} + 2\tau. \end{aligned}$$

Obviously, the duration of a node in *wait* state T_{wait} is τ .

Because by assumption collision avoidance is enforced at each node, no node is allowed to transmit data packets continuously; therefore, the transition probabilities from *succeed* to *wait* and from *fail* to *wait* are both one.

To derive the transition probability P_{ws} from *wait* to *succeed*, we need to calculate the probability $P_{ws}(r)$ that node x successfully initiates a four-way handshake with node y at a given time slot when they are at a distance r apart. Before calculating $P_{ws}(r)$, we define $B(r)$ to be the area that is in the hearing region of node y but outside the hearing region

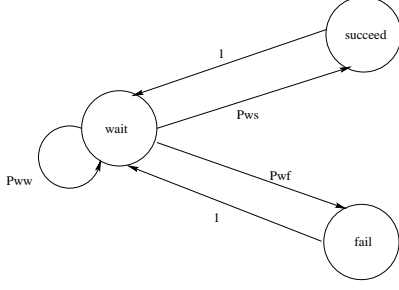


Figure 2. Markov chain model for a node

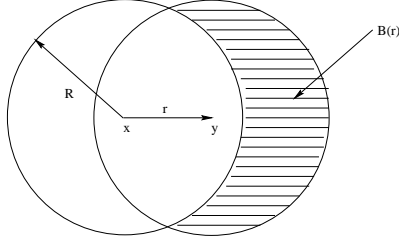


Figure 3. Illustration of "hidden" area

of node x , i.e., the interfering region "hidden" from node x as the shaded area shown in Figure 3. $B(r)$ has been shown in [13] to be:

$$B(r) = \pi R^2 - 2R^2 q\left(\frac{r}{2R}\right)$$

where $q(t) = \arccos(t) - t\sqrt{1-t^2}$.

Then $P_{ws}(r)$ can be calculated as:

$$P_{ws}(r) = P_1 \cdot P_2 \cdot P_3 \cdot P_4(r)$$

where

- $P_1 = \text{Prob.}\{x \text{ transmits in a slot}\},$
- $P_2 = \text{Prob.}\{y \text{ does not transmit in the time slot}\},$
- $P_3 = \text{Prob.}\{\text{none of the terminals within } R \text{ of } x \text{ transmits in the same slot}\},$
- $P_4(r) = \text{Prob.}\{\text{none of the terminals in } B(r) \text{ transmits for } (2l_{rts} + 1) \text{ slots} \mid r\}.$

The reason for the last term is that the vulnerable period for an RTS is only $2l_{rts} + 1$, and once the RTS is received successfully by the receiving node (which can then start sending the CTS), the probability of further collisions is assumed to be negligibly small.

Obviously, $P_1 = p'$ and $P_2 = (1 - p')$. On the other hand, P_3 can be obtained by

$$P_3 = \sum_{i=0}^{\infty} (1 - p')^i \frac{(\lambda \pi R^2)^i}{i!} e^{-\lambda \pi R^2}$$

$$\begin{aligned} &= \sum_{i=0}^{\infty} (1 - p')^i \frac{N^i}{i!} e^{-N} \\ &= e^{-p'N}. \end{aligned}$$

Similarly, the probability that none of the terminals in $B(r)$ transmits in a time slot is given by

$$\begin{aligned} p_4(r) &= \sum_{i=0}^{\infty} (1 - p')^i \frac{(\lambda B(r))^i}{i!} e^{-\lambda B(r)} \\ &= e^{-p' \lambda B(r)}. \end{aligned}$$

Hence, $P_4(r)$ can be expressed as

$$\begin{aligned} P_4(r) &= (p_4(r))^{2l_{rts}+1} \\ &= e^{-p' \lambda B(r)(2l_{rts}+1)}. \end{aligned}$$

Given that each sending node chooses any one of its neighbors with equal probability and that the average number of nodes within a region of radius r is proportional to r^2 , the probability density function of the distance r between node x and y is

$$f(r) = 2r, \quad 0 < r < 1.$$

where we have normalized r with regard to R by setting $R = 1$.

Now we can calculate P_{ws} as follows:

$$\begin{aligned} P_{ws} &= \int_0^1 2r P_{ws}(r) dr \\ &= 2p'(1 - p') e^{-p'N} \int_0^1 r e^{-p' \lambda B(r)(2l_{rts}+1)} dr \\ &= 2p'(1 - p') e^{-p'N} \cdot \int_0^1 r e^{-p'N[1-2q(r/2)/\pi](2l_{rts}+1)} dr. \end{aligned}$$

From the Markov chain shown in Figure 2, the transition probability P_{ww} that node x continues to stay in *wait* state in a slot is just $(1 - p')e^{-p'N}$, i.e., it does not initiate any transmission and there is no node around it initiating a transmission. Let π_s , π_w and π_f denote the steady-state probability of state *succeed*, *wait* and *fail*, respectively. From Figure 2, we have

$$\begin{aligned} \pi_w P_{ww} + \pi_s + \pi_f &= \pi_w \\ \pi_w P_{ww} + 1 - \pi_w &= \pi_w \\ \pi_w &= \frac{1}{2 - P_{ww}} \\ &= \frac{1}{2 - (1 - p')e^{-p'N}}. \end{aligned}$$

Therefore, the steady-state probability of state *succeed* π_s can be calculated as:

$$\pi_s = \pi_w P_{ws} = \frac{P_{ws}}{2 - (1 - p')e^{-p'N}}. \quad (2)$$

We should note that π_s is just the previous unknown quantity p_s in Equation (1). Combining Equations (1) and (2) together, we get a complex relationship between p and p' . However, given p , p' can be computed easily with numerical methods.

Accordingly, the throughput Th is:

$$\begin{aligned} Th &= \frac{\pi_s \cdot l_{data}}{\pi_w T_w + \pi_s T_s + \pi_f T_f} \\ &= l_{data} \pi_s [\tau \pi_w + (l_{rts} + l_{cts} + 2\tau)(1 - \pi_w - \pi_s) \\ &\quad + (l_{rts} + l_{cts} + l_{data} + l_{ack} + 4\tau)\pi_s]^{-1}. \end{aligned} \quad (3)$$

From the formula used to calculate throughput, we can see that π_s and π_w , from which throughput is derived, are largely dependent on p' and not on p , which is the basis for our simplification of the modeling of the channel presented earlier.

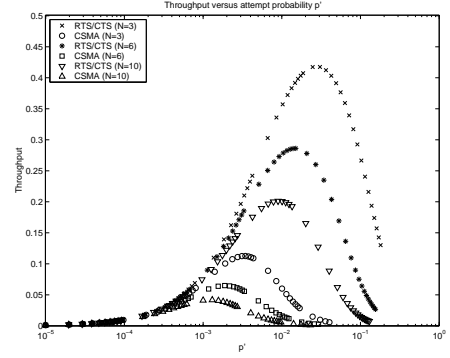
3 Numerical Results

In this section, we compare the throughput of RTS/CTS scheme with a non-persistent CSMA protocol in which there is a separate channel over which acknowledgments are sent in zero time and without collisions. The performance of the latter protocol in multi-hop networks has been analyzed by Wu and Varshney [19] and we should note that in practice, the performance of the latter protocol would be worse as both data packets and acknowledgments are transmitted in the same channel, as predicted by Tobagi and Kleinrock [15].

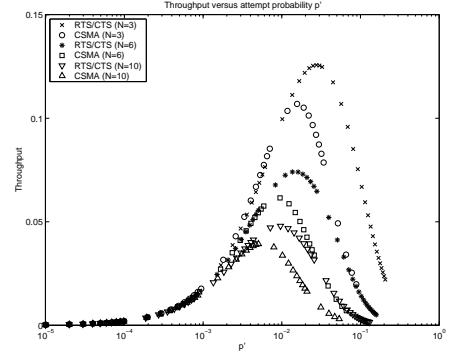
We present results for two typical cases when either relatively large data packets or relatively small data packets are sent. Let τ denote the duration of one time slot. RTS, CTS and ACK packets last 5τ . In the case of large data packets, a data packet that is much larger than the aggregate size of RTS, CTS and ACK packets. In the case of short data packets, a data packet is only slightly larger than the aggregate size of RTS, CTS and ACK packets. In the latter case, which models networks in which radios have long turn-around time and data packet transmission time is relatively short, it is uncertain whether a collision avoidance scheme should be employed due to the proportionally larger overhead.

It can be shown [16] that the maximum throughput is largely unaffected by α , and we will use $\alpha = 1$ throughout the rest of the paper. Here it should be noted that, as a side effect of not knowing the actual α that should be used, the relationship between p' and throughput may not agree with the simulations. However it will not be a problem since we are interested in the saturated throughput only.

Figure 4 compares the throughput of collision avoidance against that of CSMA with different values of N and data



(a) long data packet: $l_{data} = 100\tau$



(b) short data packet: $l_{data} = 20\tau$

Figure 4. Throughput comparison ($l_{rts} = l_{cts} = l_{ack} = 5\tau$)

packet lengths, where N is the average number of nodes that compete against one another to access the shared channel.

It is clear that the throughput of RTS/CTS scheme outperforms that of CSMA in both cases when data packet is either long or short. Even when N is as small as three, the performance of CSMA is already significantly low. This shows the importance of collision avoidance in the presence of hidden terminals. In practice, the performance margin will be larger as in the actual CSMA there is no separate idealized acknowledgment channel.

With the increase of N , the throughput of RTS/CTS scheme still degrades rapidly, even when a collision avoidance scheme is employed. A close look at Figure 4 reveals that the value of p' that achieves maximum throughput is very low. This helps to explain why maximum achievable throughput is quite limited as nodes are spending much more time on collision avoidance and backoff. This result is different from a fully-connected network, in which the

maximum throughput is largely indifferent to the number of nodes within a region [3].

Our results reveal that hidden terminals degrade the performance of collision avoidance protocols beyond the basic effect of having a longer vulnerability period for RTSs. There is one dilemma here. On the one hand, it is very difficult to get all the competing nodes around one node coordinated well by probabilistic methods such as randomized backoff. Here the competing nodes refer to both one-hop and two-hop neighbors² of the node. In actual MAC protocols, the collisions of data packets may still occur and throughput degrades with increasing number of neighbors. On the other hand, even if all the competing nodes of one node defer their access for the node, the possible spatial reuse in multi-hop networks is greatly reduced and hence the maximum achievable throughput is reduced. This dilemma leads to the scalability problem of contention-based MAC protocols that occurs much earlier than people might expect, as the throughput is already quite meager when the average of competing nodes within a region (N) is only ten.

4 Simulation Results

In this section, we investigate the performance of the popular IEEE 802.11 DFWMAC protocol via simulation experiments to validate the predictions made in the analysis.

We use GloMoSim 2.0 [20] as the network simulator. Direct sequence spread spectrum (DSSS) parameters are used throughout the simulations, which are shown in Table 1. The raw channel bit rate is 2Mbps. We use a uniform distribution to approximate the Poisson distribution used in our model, because the latter is mainly used to facilitate our derivation of analytical results and it is simply impractical to generate 2, 3, 4, ... nodes within one region, get the throughput for the individual configuration and then calculate the average. To be specific, we place nodes in concentric circles or rings. That is, given that a node's transmitting and receiving range is R and that there are on average N nodes within this circular region, we place N nodes in a circle of radius R , subject to a uniform distribution. Because there are on average 2^2N nodes within a circle of radius $2R$, we place $2^2N - N = 3N$ nodes outside the previous circle of radius R but inside the concentric circle of radius $2R$, i.e., the ring with radii R and $2R$, subject to the same uniform distribution. Then $3^2N - 2^2N = 5N$ nodes can be placed in an outer ring with radii $2R$ and $3R$.

As it is impossible to generate an infinite network in simulations, we just focus our attention on the performance of the innermost N nodes. Another reason is that it is

²Here we refer to those nodes that have at least one common neighbor with a node but are not direct neighbors of the node as the node's two-hop neighbors.

	τ	l_{rts}	l_{cts}, l_{ack}	l_{data}
actual time	21 μ sec	272 μ sec	248 μ sec	6032 μ sec
normalized	1	13	12	287

Table 2. Equivalent configuration parameters for the analytical model

more appropriate to investigate the performance of MAC schemes in a local neighborhood, rather than in the whole network as totaling and averaging performance metrics such as throughput and delay with regard to all the nodes both in the center and at the edge of a network may lead to some askew results. For example, nodes at the edge may have exceedingly high throughput due to much less contention and including them in the calculation will lead to higher than usual throughput. In our experiments, we find that nodes that are outside the concentric circles of radius $3R$ almost have no influence on the throughput of the innermost N nodes, i.e., boundary effects can be safely ignored when the circular network's radius is $3R$. Accordingly, we present only the results for a circular network of radius $3R$.

The backoff timer in the IEEE 802.11 MAC protocol is drawn from a uniform distribution whose upper bound varies according to the estimated contention level, i.e., a modified binary exponential backoff (BEB), thus p' takes on dynamic values rather than what we have assumed in the analytical model. Accordingly, we expect that the IEEE 802.11 MAC protocol will operate in a region, while our analysis gives only average performance. In addition, even in network topologies that satisfy the same uniform distribution, we can still get quite different results which will be shown later. This is also another reason that why multi-hop network model is more complex than fully-connected network model and it is worthy of more investigation.

In our simulation, each node has a constant-bit-rate (CBR) traffic generator with data packet size of 1460 bytes, and one of its neighbors is randomly chosen as the destination for each packet generated. All nodes are always backlogged. Considering the physical layer's synchronization time as well as propagation delay used in the simulation, the effective packet transmission times are shown in Table 1. For comparison purposes, we map these simulation parameters to equivalent parameters in our analytical model and they are shown in Table 2.

We run both analytical and simulation programs with $N = 3, 5$ and 8 . Though we have not tried to characterize how the performance of the IEEE 802.11 MAC protocol is distributed in the region of values taken by p' , we do have generated 50 random topologies that satisfy the uniform distribution and then get averaged transmission probability and throughput for the N nodes in the innermost circle of ra-

RTS	CTS	data	ACK	DIFS	SIFS
20-byte	14-byte	1460-byte	14-byte	50 μ sec	10 μ sec
contention window		slot time	sync. time	prop. delay	
31–1023		20 μ sec	192 μ sec	1 μ sec	

Table 1. IEEE 802.11 protocol configuration parameters

dus R for each configuration. The results are shown in Figure 5, in which the centers of rectangles are the mean values of p' and throughput and their half widths and half heights are the variance of p' and throughput respectively. These rectangles roughly describe the operating regions of IEEE 802.11 MAC protocol with the configurations we are using.

Figure 5 clearly shows that, on average, IEEE 802.11 MAC protocol cannot achieve the performance predicted in the analysis of correct collision avoidance, although the performance gap decreases when N increases. This is because that IEEE 802.11 MAC protocol cannot ensure collision-free transmission of data packets as it does not satisfy the conditions stipulated in the FAMA paper [9]. However, for some network configurations, the performance may exceed what is predicted in the analysis, especially when N is small. We reason that it results from the fairness problem in the binary exponential backoff (BEB) used in IEEE 802.11 which is aggravated in some network configurations.

The fairness problem has been investigated in the literature (e.g. [1, 2, 12]). Simply put, in BEB a node that just succeeds in sending a data packet resets its contention window to the minimum value, through which it may gain access to the channel again much earlier than other surrounding nodes. Thus, a node may monopolize the channel for a very long time during which there is no contention loss and throughput can be very high for a particular node, while other nodes suffer starvation. It can also be observed that when N increases, the variance of p' and throughput becomes smaller. Thus, the fairness problem is less severe when there are more nodes competing in a shared channel.

Given that the IEEE 802.11 MAC protocol cannot ensure that data packets are transmitted free of collisions, its throughput can deviate much from what is predicted in the analysis. To demonstrate this, we also collect statistics about the number of transmitted RTS packets that will lead to ACK timeout due to collision of data packets as well as the total number of transmitted RTS packets that can lead to either an incomplete RTS-CTS-data handshake or a successful four-way handshake. Then we calculate the ratio of these two numbers and tabulate the results in Table 3. This table clearly shows that much of the precious channel resource is wasted in sending data packets that cannot be successfully delivered.

A close observation of Figure 5 also reveals that, the gap

	$N = 3$	$N = 5$	$N = 8$
mean	0.29	0.39	0.44
std	0.17	0.10	0.06

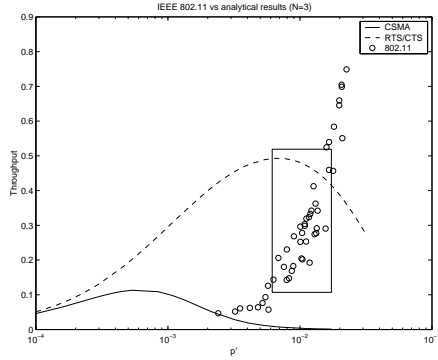
Table 3. Percentage of ACK timeout in BEB scheme

in maximum throughput between analytical and simulation results decreases when N increases. This can be explained as follows. When the number of direct competing nodes N increases, the number of indirect competing nodes (hidden terminals, $3N$ on average) also increases, which makes nodes implementing a perfect collision avoidance protocol spend much more time in deferring and backing off to coordinate with both one-hop and two-hop competing nodes to avoid collisions. Therefore, much of the gain of perfect collision avoidance is lost and possible spatial reuse is also reduced in congested area, which makes a perfect collision avoidance protocol work only marginally better than an imperfect one. This observation could not be predicted from previous analytical models or simulations focusing on fully-connected networks or networks with only a limited number of hidden terminals [3, 5, 6].

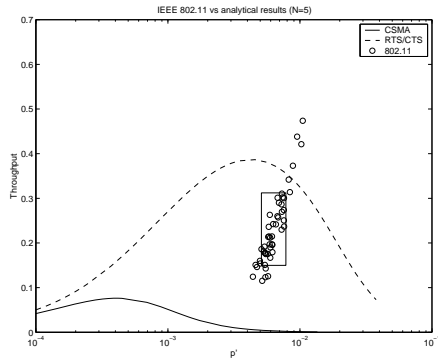
The percentage shown in Table 3 can be viewed as an *imperfectness factor* to explain the deviatory behavior of MAC protocols that do not have perfect collision avoidance. It is possible to take this factor into account in our analysis and the derivation of the theoretical throughput is not shown here due to space limit. We still show the partial results of the adjusted analysis versus the simulation results of the IEEE 802.11 protocol in Figure 6 when $N = 5$ and more results are available from the authors upon request. Figure 6 shows that the adjusted analysis is a rather good approximation of the actual performance of the IEEE 802.11 protocol though the latter has larger variation in throughput (possibly due to its inherent fairness problems).

5 Conclusion

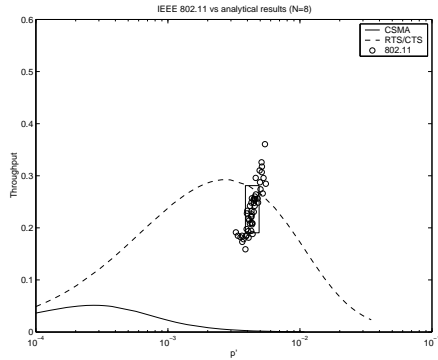
In this paper, we have used a simple model to derive the saturation throughput of MAC protocols based on an RTS-CTS-data-ACK handshake in multi-hop networks. The re-



(a) $N = 3$

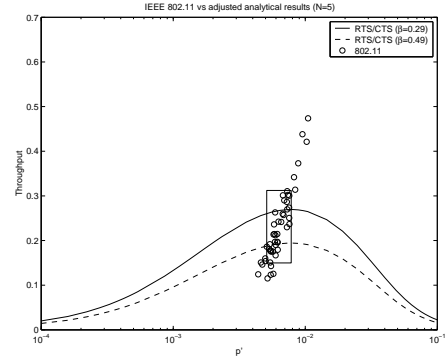


(b) $N = 5$



(c) $N = 8$

Figure 5. Performance comparison of IEEE 802.11 with analytical results



(a) $N = 5$

Figure 6. Performance comparison of IEEE 802.11 with adjusted analytical results

sults show that these protocols outperform CSMA protocols, even when the overhead of RTS/CTS exchange is rather high, thus showing the importance of correct collision avoidance in random access protocols. More importantly, it is shown that the overall performance of the sender-initiated collision avoidance scheme degrades rather rapidly when the number of competing nodes allowed within a region increases, in contrast to the case of fully-connected networks and networks with limited hidden terminals reported in the literature [3, 5, 6], where throughput remains almost the same for a large number of nodes. The significance of the analysis is that the scalability problem of contention-based collision-avoidance MAC protocols looms much earlier than people might expect. Simulation experiments with the IEEE 802.11 MAC protocol validate these observations and show that the IEEE 802.11 MAC protocol can suffer severe degradation in throughput due to its inability to avoid collisions between data packets and other packets even when the number of competing nodes in a region is small. However, when the number of competing nodes in a region increases, the performance gap is smaller as perfect collision avoidance protocols also begins to suffer from exceedingly long waiting time.

References

- [1] B. Bensaou, Y. Wang, and C. C. Ko. Fair Medium Access in 802.11 Based Wireless Ad-Hoc Networks. In *First Annual IEEE & ACM International Workshop on Mobile Ad Hoc Networking and Computing*, Aug. 2000.
- [2] V. Bharghavan, A. Demers, S. Shenker, and L. Zhang. MACAW: A Media Access Protocol for Wireless LANs. In *Proc. of ACM SIGCOMM '94*, 1994.

- [3] G. Bianchi. Performance Analysis of the IEEE 802.11 Distributed Coordination Function. *IEEE Journal on Selected Areas in Communications*, 18(3):535–547, Mar. 2000.
- [4] G. Bianchi, L. Fratta, and M. Oliveri. Performance Evaluation and Enhancement of the CSMA/CA MAC Protocol for 802.11 Wireless LANs. In *Proc. of PIMRC '96*, pages 392–396, 1996.
- [5] F. Cali, M. Conti, and E. Gregori. Dynamic Tuning of the IEEE 802.11 Protocol to Achieve a Theoretical Throughput Limit. *IEEE/ACM Transactions on Networking*, 8(6):785–799, Dec. 2000.
- [6] F. Cali, M. Conti, and E. Gregori. IEEE 802.11 Protocol: Design and Performance Evaluation of an Adaptive Backoff Mechanism. *IEEE Journal on Selected Areas in Communications*, 18(9):1774–1786, Sept. 2000.
- [7] H. S. Chhaya and S. Gupta. Throughput and Fairness Properties of Asynchronous Data Transfer Methods in the IEEE 802.11 MAC Protocol. In *Proc. of PIMRC '95*, 1995.
- [8] S. Chhaya and S. Gupta. Performance of Asynchronous Data Transfer Methods of IEEE 802.11 MAC Protocol. *IEEE Pers. Comm.*, 3(5), Oct. 1996.
- [9] C. L. Fullmer and J. J. Garcia-Luna-Aceves. Solutions to Hidden Terminal Problems in Wireless Networks. In *Proc. of ACM SIGCOMM '97*, 1997.
- [10] J. J. Garcia-Luna-Aceves and A. Tzamaloukas. Reversing The Collision-Avoidance Handshake in Wireless Networks. In *Proc. of ACM/IEEE Mobicom 1999*, Seattle, WA, U.S., 8 1999.
- [11] IEEE Computer Society LAN MAN Standards Committee, editor. *IEEE Standard for Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*. IEEE Std 802.11-1997. The Institute of Electrical and Electronics Engineers, New York, 1997.
- [12] T. Ozugur, M. Naghshineh, P. Kermani, and J. A. Copeland. Fair Media Access for Wireless LANs. In *Proc. of IEEE GLOBECOM '99*, Dec. 1999.
- [13] H. Takagi and L. Kleinrock. Optimal Transmission Range for Randomly Distributed Packet Radio Terminals. *IEEE Transactions on Communications*, 32(3):246–57, 1984.
- [14] F. A. Tobagi and L. Kleinrock. Packet Switching in Radio Channels: Part II - the Hidden Terminal Problem in Carrier Sense Multiple-access Modes and the Busy-tone Solution. *IEEE Trans. on Communications*, 23(12):1417–1433, 1975.
- [15] F. A. Tobagi and L. Kleinrock. The Effect of Acknowledgment Traffic on the Capacity of Packet-Switched Radio Channels. *IEEE Trans. on Communications*, 26(6):815–826, 1978.
- [16] Y. Wang and J. J. Garcia-Luna-Aceves. Performance of Collision Avoidance Protocols in Single-Channel Ad Hoc Networks. In *Proc. of IEEE ICNP 2002*, Paris, France, Nov. 2002.
- [17] J. Weinmiller, M. Schläger, A. Festag, and A. Wolisz. Performance Study of Access Control in Wireless LANs – IEEE 802.11 DFWMAC and ETSI RES 10 Hiperlan. *ACM/Baltzer Mobile Networks and Applications*, 2(1):55–67, 1997.
- [18] I. Widjaja, B. P. Crow, L. G. Kim, and P. T. Sakai. Investigation of the IEEE 802.11 Medium Access Control (MAC) Sublayer Functions. *Proc. of IEEE INFOCOM '97*, 1:126–133, 1997.
- [19] L. Wu and P. Varshney. Performance Analysis of CSMA and BTMA Protocols in Multihop Networks (I). Single Channel Case. *Information Sciences, Elsevier Sciences Inc.*, 120:159–77, 1999.
- [20] X. Zeng, R. Bagrodia, and M. Gerla. GloMoSim: a Library for Parallel Simulation of Large-scale Wireless Networks. In *Proc. of the 12th Workshop on Parallel and Distributed Simulations*, May 1998.